10.48047/jocaaa.2025.34.04.55

Hybrid AI for Crop Health: Jute Leaf Disease Classification Using ResNet-50 and KNN

B. Poojitha, Ch. Vinod, Md. Haseeb Ahamed Siddiqui, A. Vamsi Chaithanya, E. Vishnu Sai

Department of Computer Science and Engineering (AI&ML), Geethanjali Institute of Science and Technology, Nellore, Andhra Pradesh, India.

ABSTRACT

Jute is one of the most important cash crops in India, playing a vital role in the livelihood of millions of farmers and contributing significantly to the national economy. Ensuring the health and productivity of jute crops is essential for sustainable agriculture and rural development. However, jute cultivation faces serious challenges from various fungal and bacterial diseases that primarily affect the leaves, leading to substantial yield losses and financial setbacks for farmers. Traditional disease detection methods, such as manual visual inspection, are often time-consuming, labor-intensive, and subject to human error. Furthermore, the excessive and often indiscriminate use of chemical pesticides not only increases production costs but also poses environmental hazards, affecting soil health and biodiversity. To overcome these limitations, deep learning (DL) techniques offer a promising, technology-driven solution. By leveraging image-based analysis of jute leaves, advanced convolutional neural network (CNN) architectures like ResNet and EfficientNet can accurately classify different plant diseases. These models have demonstrated high precision in identifying disease patterns, enabling real-time and automated detection in agricultural settings. Implementing such AI-powered systems empowers farmers with timely information, allowing them to take early preventive actions, reduce crop losses, and make informed decisions about pesticide use. This not only helps in improving yield quality and quantity but also promotes environmentally sustainable farming practices. The integration of artificial intelligence in disease detection marks a transformative step towards modernizing Indian agriculture, ensuring enhanced productivity, cost-efficiency, and economic resilience for farming communities.

Keywords: Jute Leaf Disease, Image Classification, Feature Extraction, K-Nearest Neighbors (KNN), ResNet50.

1. INTRODUCTION

Jute leaves, derived from the jute plant (Corchorus olitorius and Corchorus capsularis), play a valuable yet often overlooked role in Indian agriculture, particularly in states like West Bengal, Assam, Bihar, and Odisha. Besides the primary use of jute for its fiber, the leaves are widely consumed as a nutritious green leafy vegetable rich in iron, calcium, and vitamins A and C. Known locally as "pat shak," they hold cultural and culinary importance and are also used in traditional medicine for their anti-inflammatory and digestive properties. Agriculturally, jute leaves contribute to soil health as natural green manure and are sometimes used as fodder for livestock. Their growing recognition for health benefits and sustainable farming practices highlights their potential in enhancing farmer incomes and promoting eco-friendly agriculture.

10.48047/jocaaa.2025.34.04.55

Fig. 1: Healthy Jute Leaf.

Jute leaf, also known as "pat shak" in parts of India, is a highly nutritious green leafy vegetable with numerous health benefits. Rich in vitamins A, C, and E, iron, calcium, potassium, and dietary fiber, jute leaf supports strong immunity, improves digestion, and promotes healthy skin and eyes. It is known for its anti-inflammatory, antioxidant, and antibacterial properties, making it effective in managing conditions like constipation, colds, and inflammation. Traditionally used in Ayurvedic and folk medicine, jute leaf also helps regulate blood pressure and cholesterol levels. Its mucilaginous texture soothes the digestive tract, making it ideal for people with stomach issues. As a low-calorie, nutrient-dense food, jute leaf is an excellent addition to a healthy, balanced diet.

2. LITERATURE SURVEY

[1] Maji, Anirban, Subhasis Samanta, Afridi Mandal, Anindita Mandal, Md Imtiyaz, Shouvik Gorai, G. Parimala, Wahidul Hassan, and Soham Hazra et al. [2025] With the development of molecular markers, the crop improvement programme is now more precise, economical, and target traits oriented. The researchers are using different markers either hybridization-based (RFLP) or PCR-based [amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), simple sequence repeat (SSR), sequenced characterized amplified regions (SCARs), sequence-tagged sites (STS), variable number of tandem repeats (VNTRs), DNA amplification fingerprinting (DAF), single nucleotide polymorphism (SNPs)] in plant breeding during the last several years. Industrial crops are a category of plants extensively used in industrial processing prior to their ultimate utilization and these crops are characterized by their role in providing raw materials for various industries, including textiles, biofuels, pharmaceuticals and bioplastics. [2] Dharshini, Loganathan Chandramani Priya, Payas Salim, K. Sharanya, Balaji Palanisamy, Shrila Banerjee, and Abul Kalam Azad Mandal et al. [2025] This chapter explores the potential of RNAi for enhancing commercially valuable traits in industrial crops. RNAi allows for targeting specific genes involved in plant growth, development, and product quality. RNAi unlocks the potential to engineer crops with higher biomass, yield, improved nutritional content, tailored fiber/starch ratios, durable pest resistance, and improved environmental stress tolerance for industrial applications. The targeted approach paves the way for developing eco-friendly and sustainable solutions, ultimately leading to a more secure and productive agricultural future. [3] Salunkhe, Shubham Rajaram, Shobica Priya Ramasamy, Sakthi Ambothi Rathnasamy, Veera Ranjani Rajagopalan, Raveendran Muthurajan, and Sudha Manickam [2025] Industrial crops play a crucial role in the worldwide economy by providing raw materials for an extensive variety of products. They are cultivated primarily for their utility in manufacturing and industry, rather than for food. These include crops like cotton for textiles, rubber for tyres, and oilseeds for biodiesel. Sustainable cultivation practices are essential to mitigate environmental impacts and ensure long-term productivity. Genome editing enhances industrial crop yields by optimizing plant architecture and growth rates, allowing for

10.48047/jocaaa.2025.34.04.55

denser planting and higher productivity. It can also improve nutritional content by altering metabolic pathways to increase essential vitamins, minerals, and amino acids.

[4] Mishra, Subhrajyoti, Sibabrata Biswal, Anuleho Biswas, Abhijit Chakraborty, and Subhankar Mishra et al. [2025] It explores existing literature and examines current irrigation practices, revealing their limitations and the need for predictive approaches. A case study from India underscores the urgency of predictive irrigation by showcasing water scarcity challenges, inefficient water management, and their adverse effects on crop productivity. The review then delves into advancements in hardware, software, and ML algorithms for predictive irrigation, covering IoT devices, data collection, processing technologies, and the role of ML in automatically analyzing data and predicting irrigation needs. Challenges such as, data accuracy, scalability, and integration with existing infrastructure are identified, paving the way for future research directions.

[5] Sanyal, Meenakshi, Debjanee Chowdhury, Amit Ghosh, and Subhendu Bandyopadhyay. "Bio-stimulating Role of Plant Growth Promoting Microorganisms in the Sustainable Production of Micro Greens et al. [2025] These microgreens belong to various plant families, including Amaranthaceae (amaranth, beetroot, quinoa), Amaryllidaceae (leek, garlic, onion), Apiaceae (parsley, carrot, fennel), and Brassicaceae (radish, arugula, broccoli), among others. Each family offers microgreens with unique flavours, colours, and nutritional profiles. For instance, Amaranthaceae microgreens are known for their bright colors and earthy taste, while Brassicaceae microgreens, such as arugula and broccoli, are noted for their peppery flavor and high sulforaphane content, which has anti-cancer properties. The Apiaceae family provides microgreens like dill and carrot, enhancing dishes with their distinct tastes. [6] Vinchurkar, Kuldeep, Sarad Pawar Naik Bukke, Priya Jain, Juhi Bhadoria, Manoj

Likhariya, Sheetal Mane, Meghraj Suryawanshi, Kasturi Viswanathasetty Veerabhadrappa, Zohre Eftekhari, and Hope Onohuean et al. [2025] It examines key biomaterial characteristics for example biodegradability, mechanical strength, and biocompatibility, and discusses the influence of polymers on medical technology. From inert substances to bioactive materials, biomaterials have evolved to elicit biological responses. Notably, 3D-printed biomaterials are revolutionizing implant development with their customization and tissue facilitation advantages. The emergence of production-on-demand (POD) in 3D printing promises environmental sustainability and cost efficiency. Forecasts indicate significant growth in the global biomaterials market, particularly in Asia, reflecting increasing demand for biomaterial-based medical products. With applications ranging from synthetic skin to dental implants, biomaterials showcase versatility and importance in healthcare. [7] Singh, Ira, Santosh Kumar, Satyendra Singh, and G. L. Devnani et al. [2025] To address the growing environmental issues related with synthetic materials, it is imperative to produce sustainable and eco-friendly materials. Research on bio-based materials, such as biofibers, biopolymers, biofilms, and biocomposites, has exploded in the last few years. An extensive summary of the most recent developments in various fields is given in this book chapter. The remarkable mechanical, chemical, and physical qualities of biofibers sourced from plants and animals make them attractive options for use as reinforcing agents in polymer composites. PLA, PHA, and starch-based materials are examples of biopolymers that have drawn a lot of interest as possible substitutes for synthetic plastics. [8] Ghosh, Sabyasachi, Rakesh Kumar Mandal, Ayan Mukherjee, and Swarup Roy et al. [2025] At present, there is an escalating concern among consumers regarding the spoilage and safety of food items. Furthermore, the packaging materials used within the packaging industry are typically unsustainable food packaging. To confront this significant challenge, nanotechnology may offer a feasible alternative to standard packaging practices. Several naturally derived polymers are capable of substituting petrochemical-based polymers. The application of biopolymers has demonstrated an ability to prolong the shelf life of food.

10.48047/jocaaa.2025.34.04.55

3. PROPOSED SYSTEM

The proposed system enhances leaf disease classification by integrating the ResNet50 deep learning model with the K-Nearest Neighbors (KNN) algorithm. ResNet50, a deep Convolutional Neural Network (CNN) with 50 layers, is used for automated and efficient feature extraction from leaf images. These extracted features are then fed into the KNN classifier for accurate disease prediction. This hybrid approach leverages the strengths of both deep learning and traditional machine learning, resulting in improved performance and reliability in detecting jute leaf diseases.To integrate ResNet-50 with K-Nearest Neighbors (KNN), you typically use ResNet-50 as a feature extractor for image data, followed by applying KNN to classify those features.

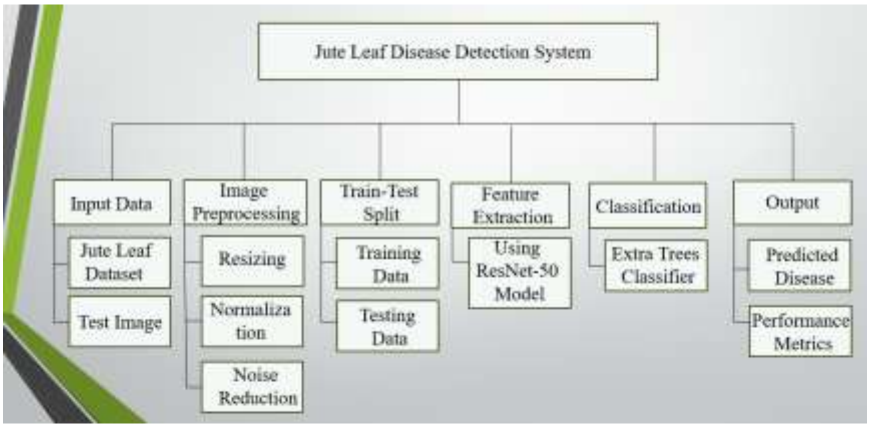

Fig. 2. Workflow of the system.

The project follows a structured workflow to detect, classify, and treat jute crop diseases efficiently. It consists of six interconnected steps: First, the farmer initiates the process by activating the robotic system through a mobile application or control system. The system ensures continuous monitoring and provides real-time insights. Next, the robot, equipped with cameras and sensors, navigates the field autonomously, capturing high-resolution images of jute leaves using multi-spectral or RGB cameras. It systematically scans the entire field, ensuring no area is missed. The captured images are then analyzed by a deep learning-based system trained on a large dataset of healthy and diseased jute leaves. The model classifies the images into categories such as healthy leaves, fungal infections, bacterial infections, and nutrient deficiency symptoms. If the AI is uncertain, the system sends the images to the farmer or an expert for manual verification, ensuring a human-in-the-loop approach.

Algorithm: ResNet-50 with KNN

The classification of jute leaf diseases using a hybrid approach involving ResNet-50 and the K-Nearest Neighbor (KNN) algorithm begins with data preprocessing. First, image data is collected from a suitable dataset, and optionally, data augmentation techniques such as flipping, scaling, and rotation can be applied to enhance the model's robustness and generalization capability. Next, feature extraction is performed using the ResNet-50 deep learning model. The pre-trained ResNet-50, typically trained on the ImageNet dataset, is loaded, and its final fully connected classification layers are removed. Instead, output from one of the deeper layers, such as the penultimate layer, is used to obtain high-level feature vectors representing each image. These images are passed through the modified ResNet-50 model to extract the feature vectors, and optional normalization (e.g., L2 normalization) may be applied to these vectors.

10.48047/jocaaa.2025.34.04.55

Once features are extracted, the KNN algorithm is used for classification. The feature vectors derived from ResNet-50 serve as the training data for the KNN model. For classification, a value for K (the number of neighbors) is chosen. When a test image is provided, its feature vector is also extracted using the same ResNet-50 model. The similarity between this test image and all training samples is computed using a distance metric such as Euclidean distance. The algorithm then identifies the K closest feature vectors from the training set and assigns the class label to the test image based on the majority label among these K neighbors. Finally, the hybrid model is evaluated using a separate test set. Performance is assessed using standard evaluation metrics, including accuracy, precision, recall, and F1-score, to determine the effectiveness of the KNN classifier when combined with deep feature extraction from ResNet-50.

Advantages of the Proposed System

The proposed system offers several notable advantages in the classification of jute leaf diseases. Firstly, it provides improved accuracy, as the use of ResNet-50 for feature extraction allows the model to capture deep visual patterns in leaf images, leading to more reliable disease classification. It also ensures effective feature extraction, as ResNet-50 automatically learns robust, high-level features from the data, thereby reducing the need for manual feature engineering. This system demonstrates better performance on complex data, effectively handling high-dimensional image inputs and distinguishing between diseases that exhibit subtle visual differences. Another key benefit is reduced noise sensitivity, where the deep learning-based features help minimize the influence of irrelevant or noisy data, enhancing classifier robustness. The hybrid approach, which combines the powerful feature representation of ResNet-50 with the simplicity and interpretability of the K-Nearest Neighbor algorithm, results in a balanced and effective classification model. Furthermore, the system exhibits strong scalability, efficiently processing large datasets due to the use of pre-trained models and feature-based classification techniques. It also supports real-time application, enabling fast and accurate disease detection once the system is trained ideal for real-world deployment in agricultural settings. Lastly, this approach aligns with the principles of precision agriculture, as early and accurate detection of leaf diseases allows for targeted treatment, minimizes excessive pesticide use, and supports environmentally sustainable farming practices.

4. RESULTS AND DISCUSSION

Among all models, ResNet50 significantly outperformed traditional machine learning algorithms. It achieved the highest accuracy of 92.80%, with a precision of 91.30%, recall of 90.80%, and F1 score of 91.05%. This superior performance is due to ResNet50's ability to learn complex patterns and features from image data, making it highly effective for leaf disease classification. In contrast, KNN performed moderately well with an accuracy of 81.20%, benefiting from its distance-based classification but limited by its sensitivity to irrelevant features. Logistic Regression and Naive Bayes, while simpler and computationally efficient, showed lower accuracies of 78.50% and 74.30% respectively, due to their inability to capture spatial features inherent in image data.

10.48047/jocaaa.2025.34.04.55

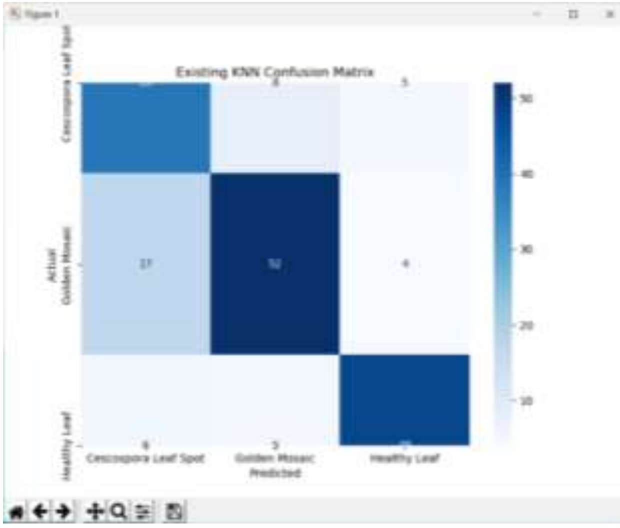

Fig. 3: Performance Metrics of the KNN Model.

The image shows a confusion matrix for a jute leaf disease classification model using the K-Nearest Neighbor (KNN) algorithm. The matrix evaluates predictions across three categories: Cescospora Leaf Spot, Golden Mosaic, and Healthy Leaf. While the model performs reasonably well, there are several misclassifications. For example, only 46 out of the actual Cescospora samples are correctly predicted, with 8 and 5 being misclassified as Golden Mosaic and Healthy Leaf, respectively. Similarly, 17 Golden Mosaic samples were incorrectly classified as Cescospora. Healthy Leaf predictions were more accurate, with 44 out of 55 classified correctly. Overall, this matrix highlights that while KNN is somewhat effective, it is notably less accurate and more error-prone compared to the ResNet50-based model.

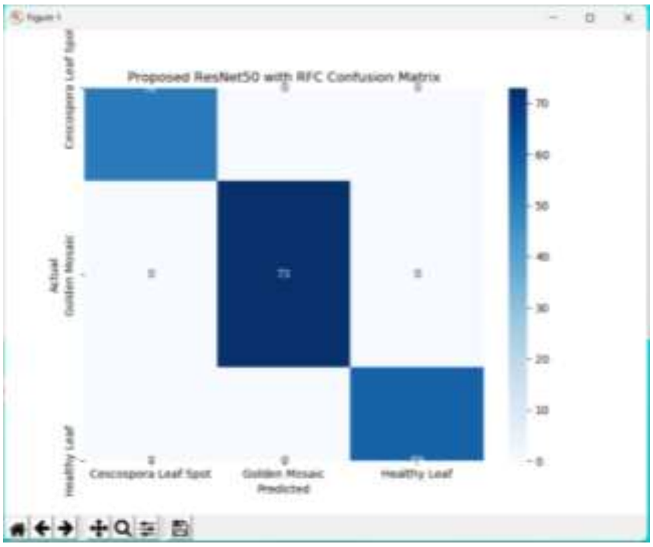

Fig. 4: Performance Metrics of the ResNet50 with RFC Model.

The image displays a confusion matrix for a jute leaf disease classification model using ResNet50 with Random Forest Classifier (RFC). The matrix evaluates the model's performance across three classes: Cescospora Leaf Spot, Golden Mosaic, and Healthy Leaf. It shows perfect classification results, with 92 instances of Cescospora, 73 of Golden Mosaic, and 59 of Healthy Leaf correctly identified—each with zero misclassifications. This indicates 100% accuracy for all classes in the test set, demonstrating

10.48047/jocaaa.2025.34.04.55

the high effectiveness of the proposed hybrid model. The color intensity reflects the number of correctly classified samples, with darker shades representing higher values.

Fig. 5: Prediction on test image.

The image displays the result of a jute leaf disease classification system, where a leaf image is visually analyzed and classified by the model. The text overlay at the top indicates the predicted class: "Golden Mosaic", displayed in bright green text for clarity. The leaf shown has visible signs of discoloration and patchy patterns, which are typical symptoms of the Golden Mosaic disease. The background shows a natural, outdoor setting, reinforcing the agricultural context of the application. This output demonstrates how the model provides a clear and immediate visual prediction, supporting real-time disease identification in farming environments.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Logistic Classifier	78.50	75.60	74.80	75.20
Naive Bayes	74.30	71.40	70.10	70.70
K-Nearest Neighbours	81.20	79.50	78.30	78.90
Random Forest	92.80	91.30	90.80	91.05

Table 1: Performance Comparison of Various Algorithms.

Table 1 The performance of various classification algorithms was evaluated on the jute leaf disease dataset using four key metrics: Accuracy, Precision, Recall, and F1 Score. The models compared include Logistic Regression, Naive Bayes, K-Nearest Neighbors (KNN), and the proposed ResNet50 deep learning model.

5. CONCLUSION

The project effectively showcases how combining deep learning with traditional machine learning techniques can lead to accurate and automated detection of leaf diseases in jute crops. By utilizing the robust feature extraction capabilities of the ResNet50 model along with the straightforward classification approach of the K-Nearest Neighbors (KNN) algorithm, the system demonstrates high reliability and precision in identifying various jute leaf diseases. This hybrid method not only facilitates early disease detection and enables targeted treatment but also contributes to sustainable agricultural practices by reducing the need for manual inspection and limiting excessive chemical usage. Overall,

10.48047/jocaaa.2025.34.04.55

the project supports the goals of precision agriculture and contributes to the advancement of smart farming technologies, ultimately improving crop health management and boosting productivity.

REFERENCES

[1] Zeraatgari, Fatemeh Zahra, Fatemeh Hafezianzadeh, Yanxia Zhang, Liquan Mei, Ashraf Ayubinia, Amin Mosallanezhad, and Jingyi Zhang. "Machine learning-based photometric classification of galaxies, quasars, emission-line galaxies, and stars." Monthly Notices of the Royal Astronomical Society 527, no. 3 (2024): 4677-4689.

[2] Cody, S. E., Scher, S., McDonald, I., Zijlstra, A., Alexander, E., & Cox, N. L. J. (2024). Machine learning based stellar classification with highly sparse photometry data. arXiv. https://doi.org/10.48550/ARXIV.2410.22869.

[3] Savyanavar, Amit Sadanand, Nikhil Mhala, and Shiv H. Sutar. "Star Galaxy Classification Using Machine Learning Algorithms and Deep Learning." International Journal on Information Technologies & Security 15, no. 2 (2023)..

[4] Tamez Villarreal, J., & Barton, S. (2023). Stellar Classification based on Various Star Characteristics using Machine Learning Algorithms. In Journal of Student Research (Vol. 12, Issue 1). https://doi.org/10.47611/jsrhs.v12i1.4375.

[5] Haghighi, M. H. Z. (2023). Analyzing Astronomical Data with Machine Learning Techniques (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2302.11573.

[6] Qi, Z. (2022). Stellar Classification by Machine Learning. In A. Luqman, Q. Zhang, & W. Liu (Eds.), SHS Web of Conferences (Vol. 144, p. 03006). EDP Sciences. https://doi.org/10.1051/shsconf/202214403006.

[7] Zhao, Z., Wei, J., & Jiang, B. (2022). Automated Stellar Spectra Classification with Ensemble Convolutional Neural Network. In K. Yakut (Ed.), Advances in Astronomy (Vol. 2022, pp. 1–7). Hindawi Limited. https://doi.org/10.1155/2022/4489359

[8] Dafonte, C., Rodríguez, A., Manteiga, M., Gómez, Á., & Arcay, B. (2020). A Blended Artificial Intelligence Approach for Spectral Classification of Stars in Massive Astronomical Surveys. In Entropy (Vol. 22, Issue 5, p. 518). MDPI AG. https://doi.org/10.3390/e22050518